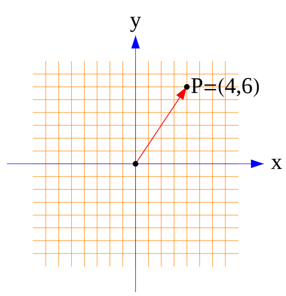
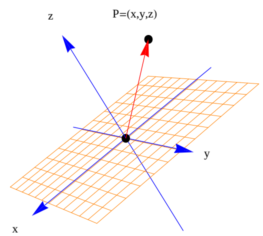
Oliver Knill, Harvard Summer School, 2009

Chapter 1. Geometry and Space

Section 1.1: Space, distance, geometrical objects

A point P on the real line is labeled by a single **coordinate** $P = x$, a point in the **plane** is fixed by 2 **coordinates** $P = (x, y)$ and a point in space is determined by three coordinates $P = (x, y, z)$. Depending on which coordinates are positive, one can divide the line, the plane or the space into 2 **half lines**, 4 **quadrants** or 8 **octants**. The point $O = 0, O = (0, 0)$ or $O = (0, 0, 0)$ is called the **origin**.





In order to get these coordinates, we need a **coordinate system**, three coordinate axes which are usually assumed to be perpendicular to each other. We call them the x-coordinate axes, the y-coordinate axes and the z-coordinate axes. The choice of the convenient coordinate system depends on the situation. On earth for example, the coordinate system is often chosen so that the z-axis points "up" and is perpendicular to the x-y plane which forms the "ground". In two dimensions, on a sheet of paper, the x-coordinate usually is chosen to point "east" and the y-coordinate to point "north". But this does not need always to be so:
In 3D computer graphics like computer games, virtual reality or ray tracing, it is custom to have the y-axis pointing up, the x-axis to the right and the z-axis in front. This is called the **"photographers coordinate system"** because the photo plate is the xy-plane, then the depth is the z-axis. Also in computer graphics, the part of the memory which is reserved for storing the z-axis is called the **"z-buffer"**. It is useful for "hidden line removal" in 2D rendering of a 3D scene: The z-axis is perpendicular to the screen with values increasing towards the viewer. Any point whose z-coordinate is smaller than the corresponding z-buffer value will be hidden behind parts which are already plotted. The photographers coordinate system is oriented differently than the usual coordinate system. You can explore this by taking the right hand, pointing the thumb in the first direction, the pointing finger into the second direction and the middle finger into the third direction.

The **Euclidean distance** between points $P = (x, y, z)$ and $Q = (a, b, c)$ is defined as

$$\boxed{d(P,Q) = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}}$$

While other distances can be defined in space, like the trivial distance $d_1(x, y) = 1$ if $x \neq y$ and $d(x, x) = 0$, or the "Manhattan distance" $d_2(P, Q) = |x - a| + |y - b| + |z - c|$, the Euclidean distance is distinguished in that it does not change under rotations or translations or combinations of those and that the sphere is a surface. These properties force the distance to be the Euclidean distance. The derivation of the distance formula uses Pythagoras theorem, which itself can be proved in a purely geometric way under the assumption that length is invariant under rotations and translations.

We usually work with a **right handed coordinate system**. The photographers coordinate system is an example of a **left handed coordinate system**. The **"right hand rule"** tells that if we use the thumb for the x-direction the index finger=y-direction and the middle finger=z-direction, then the coordinate system is "right handed". This left or right handedeness is also called **Parity**. It is relevant in biology because DNA or Proteins have a distinguished orientation. It is also important in particle physics, where "parity violation" can happen: physical laws change when we look at them in a mirror. Coordinate systems with different parity can not be rotated into each other. One would need a reflection, a "mirror" to do so.

Curves, **surfaces** and **bodies** are examples of geometrical objects which can be described using **functions of several variables**. We focus in this first lecture on spheres or circles. A **circle** of radius r centered at $P = (a, b)$ is the collection of points in the plane which have distance r from P. A **sphere** of radius ρ centered at $P = (a, b, c)$ is the collection of points in space which have the distance ρ from P. The equation of a sphere is

$$\boxed{(x-a)^2 + (y-b)^2 + (z-c)^2 = \rho^2}$$

Here is a reminder how we solve a quadratic equation. First divide the equation to have it in the form $x^2 + bx + c = 0$. We then add $(b/2)^2 - c$ on both sides. This **"completion of the square"** is due to the mathematician **Al-Khwarizmi**: gives $(x + b/2)^2 = (b/2)^2 - c$. Solving for x is the usual formula for the root of quadratic equations. We can use the completion of squares also to understand equations with several variables. For example, the equation $x^2+5x+y^2-2y+z^2 = -1$ is after completion of the square $(x+5/2)^2-25/4+(y-1)^2-1+z^2 = -1$ or $(x - 5/2)^2 + (y - 1)^2 + z^2 = (5/2)^2$. We see a sphere **center** $(5/2, 1, 0)$ and **radius** $5/2$.
In an appendix to "Geometry" to his "Discours de la méthode", René Descartes (1596-1650) promoted the idea that algebra could be used as a general method to solve geometric problems. Even so Descartes mostly algebraized ruler-and compass constructions, in honor of Descartes, the rectangular coordinate system is called the **Cartesian coordinate system**. In its current form, Cartesian geometry is due as much to Descartes own contemporaries and successors as to himself (Davis-Hersh 1986). Some anecdote: "In 1649, Queen Christina of Sweden persuaded Descartes to go to Stockholm. Because the Queen wanted to "draw tangents" at 5 AM, Descartes broke the usual habit of getting up at 11 AM. But after only a few months in the cold northern climate, walking to the palace early at 4 AM in the morning, Descartes died of pneumonia. Others think, that he might have been poisoned because he had too much influence on Christina. See the book "Descartes Secret Notebook" by Amir D. Aczel.

We focus to 2 and 3 dimensions in this course. But what about higher dimensions? We have seen that in two dimensions, the coordinate axis $x = 0$, $y = 0$ divides the plane into 4 regions called **quadrants**. Similarly, the coordinate planes $x = 0$, $y = 0$ and $z = 0$ divide the space into 8 regions which are called **octants**. This could be continued into higher dimensions: how many "hyper-regions" are there in four dimensional "hyper-space" which is labeled by points with 4 coordinates (t, x, y, z)? The answer is that there are 16 hyper-regions and each of them contains one of the 16 points (x, y, z, w), where x, y, z, w are either $+1$ or -1.

Section 1.2: Vectors, dot product, projections

Two points $P = (a, b, c)$, $Q = (x, y, z)$ define a **vector** $\vec{v} = \langle x - a, y - b - z - c\rangle$. It points from P to Q and we can write also $\vec{v} = \vec{PQ}$. The real numbers numbers p, q, r in a vector $\vec{v} = \langle p, q, r\rangle$ are called **components** of $\vec{v}$. Vectors can be drawn **everywhere** in space. If a vector starts at the origin $O = (0, 0, 0)$, then $\vec{v} = \langle p, q, r\rangle$ points to the point (p, q, r). One can therefore identify points $P = (a, b, c)$ with vectors $\vec{v} = \langle a, b, c\rangle$ attached at the origin. Two vectors with the same components are considered **equal** if the translate into each other that is if their components are the same.

The **sum** of two vectors is $\vec{u} + \vec{v} = \langle u_1, u_2\rangle + \langle v_1, v_2\rangle = \langle u_1 + v_1, u_2 + v_2\rangle$ a **scalar multiple** $\lambda\vec{u} = \lambda\langle u_1, u_2\rangle = \langle \lambda u_1, \lambda u_2\rangle$. The difference $\vec{u} - \vec{v}$ can best be seen as the addition of $\vec{u}$ and $(-1) \cdot \vec{v}$.

The vectors $\vec{i} = \langle 1, 0\rangle$, $\vec{j} = \langle 0, 1\rangle$ are called **standard basis vectors** in the plane. In space, one has the basis vectors $\vec{i} = \langle 1, 0, 0\rangle$, $\vec{j} = \langle 0, 1, 0\rangle$, $\vec{k} = \langle 0, 0, 1\rangle$.
Every vector $\vec{v} = \langle p, q\rangle$ in the plane can be written as $\vec{v} = p\vec{i} + q\vec{j}$. Every vector $\vec{v} = \langle p, q, r\rangle$ in space can be written as $\vec{v} = p\vec{i} + q\vec{j} + r\vec{k}$.

Vectors are abundant in applications. They appear for example in mechanics: if $\vec{r}(t) = \langle f(t), g(t)\rangle$ is a point in the plane which depends on time t, then $\vec{v} = \langle f'(t), g'(t)\rangle$ will be called the **velocity vector** at $\vec{r}(t)$. Some problems in statics involve the determination of a forces acting on objects and forces are represented as vectors. In particular, electromagnetic or gravitational fields or velocity fields in fluids are described by vectors. Vectors appear also in computer science: the scalable Vector Graphics is an emerging standard for the web for describing two-dimensional graphics. Objects in it are described by vectors. In quantum computation, rather than working with bits, one deals with **qbits**, which are vectors. Finally, **color** can be written as a vector $\vec{v} = \langle r, g, b\rangle$, where r is **red**, g is **green** and b is **blue** component of the color vector. An other coordinate system for color is $\vec{v} = \langle c, m, y\rangle = \langle 1 - r, 1 - g, 1 - b\rangle$, where c is **cyan**, m is **magenta** and y is **yellow**.

The addition and scalar multiplication of vectors satisfy "obvious" properties. There is no need to memorize these properties. We write $*$ here for multiplication with a scalar but usually, the multiplication sign is left out. Here is a list of properties: **commutativity** $\vec{u} + \vec{v} = \vec{v} + \vec{u}$, **associativity** $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ and $r * (s * \vec{v}) = (r * s) * \vec{v}$ as well as **distributivity** $(r + s)\vec{v} = \vec{v}(r + s)$ and $r(\vec{v} + \vec{w}) = r\vec{v} + r\vec{w}$.

The length $|\vec{v}|$ of a vector $\vec{v} = \vec{PQ}$ is defined as the distance from the beginning to the end of the vector. For example if $\vec{v} = \langle 3, 4\rangle$, then $|\vec{v}| = \sqrt{25} = 5$. More examples are $|\vec{i}| = |\vec{j}| = |\vec{k}| = 1$, $|\vec{0}| = 0$. and $|\langle 3, 4, 12\rangle| = 13$.

A vector of length 1 is called a **unit vector**. If $\vec{v} \neq \vec{0}$, then $\vec{v}/|\vec{v}|$ is a unit vector. For example, if $\vec{v} = \langle 3, 4\rangle$, then $\vec{v} = \langle 3/5, 4/5\rangle$ is a unit vector, $\vec{i}, \vec{j}, \vec{k}$ are unit vectors.

Two vectors $\vec{v}$ and $\vec{w}$ are called **parallel**, if $\vec{v} = r\vec{w}$ with some constant r.

The **dot product** of two vectors $\vec{v} = \langle a, b, c\rangle$ and $\vec{w} = \langle p, q, r\rangle$ is defined as $\vec{v} \cdot \vec{w} = ap + bq + cr$.

Different notations are used in different fields but it is all the dot product: while mathematicians write $\vec{v} \cdot \vec{w} = (\vec{v}, \vec{w})$, one can see $\langle \vec{v}|\vec{w}\rangle$ in quantum mechanics, the Einstein notation $v_i w^i$ or more generally $g_{ij} v^i w^j$ in general relativity. The dot product is also called **scalar product**, or **inner product**. Using the dot product one can express the length of $\vec{v}$ as $|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}$.

Can we express the dot product in terms of the length alone? The answer is yes: $(\vec{v}+\vec{w})\cdot(\vec{v}+\vec{w}) = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w} + 2(\vec{v} \cdot \vec{w})$ can be solved for $\vec{v} \cdot \vec{w}$.

Because $|\vec{v} - \vec{w}|^2 = (\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w}) = |\vec{v}|^2 + |\vec{w}|^2 - 2\vec{v} \cdot \vec{w}$, which is by the cos-**theorem** (which is also called Al Kashi's theorem) equal to $|\vec{v}|^2 + |\vec{w}|^2 - 2|\vec{v}| \cdot |\vec{w}| \cos(\alpha)$, with the angle α between $\vec{v}, \vec{w}$, we obtain the important cos- **formula**

$$\boxed{\vec{v} \cdot \vec{w} = |\vec{v}| \cdot |\vec{w}| \cos(\alpha)}$$

The **Cauchy-Schwartz identity** $|\vec{v} \cdot \vec{w}| \leq |\vec{v}||\vec{w}|$ follows from the cos-formula and $|\cos(\alpha)| \leq 1$. This inequality can be proven directly without referring to the cos-formula: It is enough to prove it for $|w| = 1$. Now plug in $a = x \cdot y$ into the equation $0 \leq (v - aw) \cdot (v - aw)$ to get $0 \leq (v - (v \cdot w)w) \cdot (v - (v \cdot w)w) = |v|^2 + (v \cdot w)^2 - 2(v \cdot w)^2 = |v|^2 - (v \cdot w)^2$ which means $(v \cdot w)^2 \leq |v|^2$.

The **triangle inequality** $|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$ follows from $|\vec{u} + \vec{v}|^2 = (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v}) = \vec{u}^2 + \vec{v}^2 + 2\vec{u} \cdot \vec{v} \leq \vec{u}^2 + \vec{v}^2 + 2|\vec{u} \cdot \vec{v}| \leq \vec{u}^2 + \vec{v}^2 + 2|\vec{u}| \cdot |\vec{v}| = (|\vec{u}| + |\vec{v}|)^2$.

Two vectors are called **orthogonal** or **perpendicular** if $\vec{v} \cdot \vec{w} = 0$. The zero vector $\vec{0}$ is orthogonal to any vector. For example, $\vec{v} = \langle 2, 3\rangle$ is orthogonal to $\vec{w} = \langle -3, 2\rangle$.

We can even recover **Pythagoras theorem** using our new language: Pythagoras tells that if $\vec{v}$ and $\vec{w}$ are orthogonal, then $|v - w|^2 = |v|^2 + |w|^2$.
Here is the algebraic proof $(\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w}) = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w} + 2\vec{v} \cdot \vec{w} = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w}$. Note that we have used Pythagoras to derive the cos-formula so that this looks like a "circulus vitiosus". But we can still look at this as a proof if we consider the **cos**-formula as the **definition** of angle.

The vector

$$\text{proj}_{\vec{w}}(\vec{v}) = \frac{\vec{v} \cdot \vec{w}}{|\vec{w}|^2}\vec{w}$$

is called the **projection** of $\vec{v}$ onto $\vec{w}$. The **scalar projection**

$$\frac{\vec{v} \cdot \vec{w}}{|\vec{w}|} .$$

is a signed length of the vector projection. Its absolute value is the length of the projection of $\vec{v}$ onto $\vec{w}$. The vector $\vec{b} = \vec{v} - \vec{a}$ is called the **component** of $\vec{v}$ orthogonal to the $\vec{w}$-direction.

For example, with $\vec{v} = \langle 0, -1, 1\rangle$, $\vec{w} = \langle 1, -1, 0\rangle$, $\text{P}_{\vec{w}}(\vec{v}) = \langle 1/2, -1/2, 0\rangle$. The scalar projection is $1/\sqrt{2}$.

We will use the projection to compute distances between various objects.

Section 1.3: The cross product and triple scalar product

The **cross product** of two vectors $\vec{v} = \langle v_1, v_2, v_3 \rangle$ and $\vec{w} = \langle w_1, w_2, w_3 \rangle$ is defined as the vector

$$\boxed{\vec{v} \times \vec{w} = \langle v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1 \rangle .}$$

To remember it, write it as a "determinant":

$$\begin{bmatrix} i & j & k \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} = \begin{bmatrix} i & & \\ & v_2 & v_3 \\ & w_2 & w_3 \end{bmatrix} - \begin{bmatrix} & j & \\ v_1 & & v_3 \\ w_1 & & w_3 \end{bmatrix} + \begin{bmatrix} & & k \\ v_1 & v_2 & \\ w_1 & w_2 & \end{bmatrix}$$

which is $\vec{i}(v_2w_3 - v_3w_2) - \vec{j}(v_1w_3 - v_3w_1) + \vec{k}(v_1w_2 - v_2w_1)$.

The cross product $\vec{v} \times \vec{w}$ is orthogonal to $\vec{v}$ and orthogonal to $\vec{w}$. This can be checked directly. We just have to verify for example that $\vec{v} \cdot (\vec{v} \times \vec{w}) = 0$. An important formula is

$$\boxed{|\vec{v} \times \vec{w}| = |\vec{v}||\vec{w}| \sin(\alpha)}$$

Proof: Verify first the **Lagrange's identity** $|\vec{v} \times \vec{w}|^2 = |\vec{v}|^2|\vec{w}|^2 - (\vec{v} \cdot \vec{w})^2$ by direct computation. Now, $|\vec{v} \cdot \vec{w}| = |\vec{v}||\vec{w}| \cos(\alpha)$.

The length $|\vec{v} \times \vec{w}|$ is the area of the parallelogram spanned by $\vec{v}$ and $\vec{w}$.

Proof. Because $|\vec{w}| \sin(\alpha)$ is the height of the parallelogram with base length $|\vec{v}|$, the area is $|\vec{v}||\vec{w}| \sin(\alpha)$ which is by the above formula equal to $|\vec{v} \times \vec{w}|$.

For example, if $\vec{v} = (a, 0, 0)$ and $\vec{w} = (b\cos(\alpha), b\sin(\alpha), 0)$, then $\vec{v} \times \vec{w} = (0, 0, ab\sin(\alpha))$ which has length $|ab\sin(\alpha)|$.

We see that $\vec{v} \times \vec{w}$ is zero if $\vec{v}$ and $\vec{w}$ are **parallel**.

The vectors $\vec{v}, \vec{w}$ and $\vec{v} \times \vec{w}$ form a **right handed coordinate system**. The right hand rule is: if the first vector $\vec{v}$ is the thumb, the second vector $\vec{w}$ is the pointing finger then $\vec{v} \times \vec{w}$ is the third middle finger of the right hand. For example, the vectors $\vec{i}, \vec{j}, \vec{i} \times \vec{j} = \vec{k}$ form a right handed coordinate system.

The scalar $[\vec{u}, \vec{v}, \vec{w}] = \vec{u} \cdot (\vec{v} \times \vec{w})$ is called the **triple scalar product** of $\vec{u}, \vec{v}, \vec{w}$. It is a scalar.

The absolute value of $[\vec{u}, \vec{v}, \vec{w}]$ is the volume of the parallelepiped spanned by $\vec{u}, \vec{v}, \vec{w}$ because $h = \vec{u} \cdot \vec{n}/|\vec{n}|$ is the height of the parallelepiped if $\vec{n} = (\vec{v} \times \vec{w})$ is a normal vector to the ground parallelogram which has area $A = |\vec{n}| = |\vec{v} \times \vec{w}|$. The volume of the parallelepiped is $hA = (\vec{u} \cdot \vec{n}/|\vec{n}|)|\vec{v} \times \vec{w}|$ which simplifies to $(\vec{u} \cdot \vec{n} = |(\vec{u} \cdot (\vec{v} \times \vec{w})|$ which is indeed the absolute value of the triple scalar product.

For example, to find the volume of the parallelepiped with corners $O = (1, 1, 0), P = (2, 3, 1), Q = (4, 3, 1), R = (1, 4, 1)$, we first see that it is spanned by the vectors $\vec{u} = \langle 1, 2, 1 \rangle$, $\vec{v} = \langle 3, 2, 1 \rangle$, and $\vec{w} = \langle 0, 3, 2 \rangle$. We get $\vec{v} \times \vec{w} = \langle 1, -6, 9 \rangle$ and $\vec{u} \cdot (\vec{v} \times \vec{w}) = -2$. The volume is 2.

Section 1.4: Lines, planes and distances

A point P and a vector $\vec{v}$ define a line L. It is the set of points $L = \{\vec{OP} + t\vec{v}$, where t is a real number $\}$. The line contains the point P and points into the direction of $\vec{v}$. Every vector you draw in the line is parallel to $\vec{v}$.
Sometimes a **parameter interval** $[a, b]$ is given and t assumed to be in that interval. In that case, we have a **line segment** which connects $\vec{r}(a)$ with $\vec{r}(b)$.
Assume we want to get the line through the points $P = (1, 1, 2)$ and $Q = (2, 4, 6)$, we form the vector $\vec{v} = \vec{PQ} = \langle 2, 4, 6 \rangle$ and get $L = \{\langle x, y, z \rangle = \langle 1, 1, 2 \rangle + t\langle 2, 4, 6 \rangle; \}$. This can be written also as $\vec{r}(t) = \langle 1 + 2t, 1 + 4t, 2 + 6t \rangle$. This description is called the **parametric equation** for the line. The parameter t can be thought of as "time".

If we write $\langle x, y, z \rangle = \langle 1, 1, 2 \rangle + t\langle 2, 4, 6 \rangle$ as a collection of equations $x = 1 + 2t, y = 1 + 4t, z = 2 + 6t$ and solve the first equation for t and plug it into the other equations, we get $y = 1 + (2x - 2), z = 2 + 3(2x - 2)$. The line was described as

$$L = \{(x, y, z) \mid y = 2x - 1, z = 6x - 4\ \} .$$

More generally, the line $\vec{r} = \vec{OP} + t\vec{v}$ with $P = (p, q, r)$ and $\vec{v} = \langle a, b, c \rangle$ satisfies the **symmetric equations**

$$\boxed{\frac{x-p}{a} = \frac{y-q}{b} = \frac{z-r}{c}}.$$

Every of these expressions is equal to t.

For example, to find the symmetric equations for the line through the two points $P = (0, 1, 1)$ and $Q = (2, 3, 4)$, we first form the parametric equations are $\langle x, y, z \rangle = \langle 0, 1, 1 \rangle + t\langle 2, 2, 3 \rangle$ or $x = 2t, y = 1 + 2t, z = 1 + 3t$. Solving each equation for t gives the symmetric equation $x/2 = (y - 1)/2 = (z - 1)/3$.

A point P and two vectors $\vec{v}, \vec{w}$ define a **plane** Σ. It can be defined as the set of points $\Sigma = \{\vec{OP} + t\vec{v} + s\vec{w}$, where t, s are real numbers $\}$.

An example is $\Sigma = \{\langle x, y, z \rangle = \langle 1, 1, 2 \rangle + t\langle 2, 4, 6 \rangle + s\langle 1, 0, -1 \rangle\ \}$. This is called the **parametric description** of a plane.

If a plane contains the two vectors $\vec{v}$ and $\vec{w}$, then the vector $\vec{n} = \vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$. Because also the vector $\vec{PQ} = \vec{OQ} - \vec{OP}$ is perpendicular to $\vec{n}$, we have $(Q - P) \cdot \vec{n} = 0$. With $Q = (x_0, y_0, z_0)$, $P = (x, y, z)$, and $\vec{n} = \langle a, b, c \rangle$, this means $ax + by + cz = ax_0by_0 + cz_0 = d$. The plane is therefore described by a single equation $ax + by + cz = d$.

A typical problem is to find the equation of a plane which contains the three points, like $P = (-1, -1, 1), Q = (0, 1, 1), R = (1, 1, 3)$.

To solve this, note that the plane contains the two vectors $\vec{v} = \langle 1, 2, 0 \rangle$ and $\vec{w} = \langle 2, 2, 2 \rangle$. We have $\vec{n} = \langle 4, -2, -2 \rangle$ and the equation is $4x - 2y - 2z = d$. The constant d is obtained by plugging in the coordinates of a point to the left. In our case, it is $4x - 2y - 2z = -4$.

The **angle between the two planes** $ax + by + cz = d$ and $ex + fy + gz = h$ is $\arccos(\vec{n}_1 \cdot \vec{n}/(|\vec{n}|\vec{m}|)$, where $\vec{n} = \langle a, b, c\rangle$ and $\vec{m} = \langle e, f, g\rangle$. Alternatively, it is $\arcsin(|\vec{n} \times \vec{m}|/(|\vec{n}||\vec{m}|))$.

To find the **line of intersection** of two non-parallel planes $ax+by+cz = d$ and $ex+fy+gz = h$, first find a point P which is in the intersection. Then $\vec{r}(t) = \vec{OP} + t(\vec{n} \times \vec{m})$ is the line, we were looking for.

Finally note that lines in the plane can also be parametrized in the same way: $\vec{r}(t) = \vec{OQ} + t\vec{v}$. Eliminating t gives a single equation of the form $ax+by = d$. For example, $\langle x, y\rangle = \langle 1, 2\rangle + t\langle 3, 4\rangle$ is equivalent to $x = 1 + 3t, y = 2 + 4t$ and so $4x - 3y = -2$.
To the end of this section and this chapter, lets look at some distance formulas:

1) If P is a point and $\Sigma : \vec{n} \cdot \vec{x} = d$ is a plane containing a point Q, then

$$d(P, \Sigma) = \frac{|\vec{PQ} \cdot \vec{n}|}{|\vec{n}|}$$

is the distance between P and the plane. Proof: use the angle formula in the denominator. For example, to find the distance from $P = (7, 1, 4)$ to $\Sigma : 2x + 4y + 5z = 9$, we find first a a point $Q = (0, 1, 1)$ on the plane. Then compute

$$d(P, \Sigma) = \frac{|\langle -7, 0, -3\rangle \cdot \langle 2, 4, 5\rangle|}{|\langle 2, 4, 5\rangle|} = \frac{29}{\sqrt{45}} .$$

2) If P is a point in space and L is the line $\vec{r}(t) = Q + t\vec{u}$, then

$$d(P, L) = \frac{|(\vec{PQ}) \times \vec{u}|}{|\vec{u}|}$$

is the distance between P and the line L. Proof: the area divided by base length is height of parallelogram. For example, to compute the distance from $P = (2, 3, 1)$ to the line $\vec{r}(t) = (1, 1, 2) + t(5, 0, 1)$, compute

$$d(P, L) = \frac{|\langle -1, -2, 1\rangle \times \langle 5, 0, 1\rangle|}{\langle 5, 0, 1\rangle} = \frac{|\langle -2, 6, 10\rangle|}{\sqrt{26}} = \frac{\sqrt{140}}{\sqrt{26}} .$$

3) If L is the line $\vec{r}(t) = Q + t\vec{u}$ and M is the line $\vec{s}(t) = P + t\vec{v}$, then

$$d(L, M) = \frac{|(\vec{PQ}) \cdot (\vec{u} \times \vec{v})|}{|\vec{u} \times \vec{v}|}$$

is the distance between the two lines L and M. Proof: the distance is the length of the vector projection of $\vec{PQ}$ onto $\vec{u} \times \vec{v}$ which is normal to both lines. For example, to compute the distance between $\vec{r}(t) = (2, 1, 4) + t(-1, 1, 0)$ and M is the line $\vec{s}(t) = (-1, 0, 2) + t(5, 1, 2)$ form the cross product of $\langle -1, 1, 0\rangle$ and $\langle 5, 1, 2\rangle$ is $\langle 2, 2, -6\rangle$. The distance between these two lines is

$$d(L, M) = \frac{|(3, 1, 2) \cdot (2, 2, -6)|}{|\langle 2, 2, -6\rangle|} = \frac{4}{\sqrt{44}} .$$

4) To get the distance between two planes $\vec{n} \cdot \vec{x} = d$ and $\vec{n} \cdot \vec{x} = e$, then their distance is

$$\frac{|e - d|}{|\vec{n}|}$$

Non-parallel planes have distance 0. Proof: use the distance formula between point and plane. For example, $5x + 4y + 3z = 8$ and $10x + 8y + 6z = 2$ have the distance

$$\frac{|8 - 1|}{|\langle 5, 4, 3\rangle|} = \frac{7}{\sqrt{50}} .$$

Lets mention a distance problem which has a great deal of application and motivates the material of the next week:
The global positioning system GPS uses the fact that a receiver can get the difference of distances to two satellites. Each GPS satellite sends periodically signals which are triggered by an atomic clock. While the distance to each satellite is not known, the difference from the distances to two satellites can be determined from the time delay of the two signals. This clever trick has the consequence that the receiver does not need to contain an atomic clock itself. To understand this better, we need to know about functions of three variables and surfaces.